上海大学出版社
2005年上海大学博士学位论文 91

基于H.264/AVC的可伸缩视频编码及其相关技术研究

• 作 者：李志刚
• 专 业：通信与信息系统
• 导 师：张兆扬

G643/145
001280743

上海大学出版社

2005年上海大学博士学位论文 91

基于H.264/AVC的可伸缩视频编码及其相关技术研究

- 作 者：李志刚
- 专 业：通信与信息系统
- 导 师：张兆扬

贵阳学院图书馆

图书在版编目(CIP)数据

2005年上海大学博士学位论文.第2辑/博士论文编辑部编.—上海：上海大学出版社，2009.6
ISBN 978-7-81118-367-2

Ⅰ.2… Ⅱ.博… Ⅲ.博士—学位论文—汇编—上海市—2005 Ⅳ.G643.8

中国版本图书馆CIP数据核字(2008)第180878号

2005年上海大学博士学位论文

——第2辑

上海大学出版社出版发行
(上海市上大路99号 邮政编码200444)
(http://www.shangdapress.com 发行热线66135110)
出版人：姚铁军

*

南京展望文化发展有限公司排版
上海华业装潢印刷厂印刷 各地新华书店经销
开本890×1240 1/32 印张274.25 字数7641千
2009年6月第1版 2009年6月第1次印刷
印数：1～400
ISBN 978-7-81118-367-2/G·490 定价：980.00元(49册)

Shanghai University Doctoral Dissertation (2005)

Study on the Scalable Video Coding Based on H. 264/ AVC Standard and Its Related Technologies

Candidate: Li Zhigang
Major: Communication and Information System
Supervisor: Zhang Zhaoyang

Shanghai University Press
• Shanghai •

上海大学

本论文经答辩委员会全体委员审查，确认符合上海大学博士学位论文质量要求。

答辩委员会名单：

主任：	余松煜	教授，上海交大图像通信所	200030
委员：	王朔中	教授，上海大学电子信息工程系	200072
	王国中	教授，上广电中央研究院	200333
	王治钢	研究员，航天局 809 研究所	200050
	方　勇	教授，上海大学电子信息工程系	200072
导师：	张兆扬	教授，上海大学	200072

此为试读，需要完整PDF请访问：www.ertongbook.com

评阅人名单：

余松煜	教授，上海交大图像通信所	200030
王朔中	教授，上海大学电子信息工程系	200072
王汝笠	研究员，中科院技术物理所	200083

评议人名单：

张立明	教授，复旦大学电子工程系	200433
严壮志	教授，上海大学生物医学工程研究所	200072
周源华	教授，上海交大图像通信所	200030
翁默颖	教授，华东师大电子科学技术系	200062

答辩委员会对论文的评语

李志刚同学的博士学位论文《基于 H.264/AVC 的可伸缩视频编码及其相关技术研究》是国家自然科学基金重点项目(60332030)中的重要内容之一。H.264 和可伸缩编码是国内外关注的研究热点,前景广阔,论文选题有重要的理论意义和实用价值,且在以下几个方面取得创新性研究成果:

1. 在 H.264 算法优化和指令集优化的基础上实现了 PC 平台上的实时 FGS 编解码。

2. 提出一种基于 H.264 的空间/时间/SNR FGS 方案。该方案在扩大了可伸缩范围的同时,较明显地提高了 FGS 编码的性能,并对它的码率控制策略作了较深入的研究。

3. 提出一种将 H.264 流切换和 FGS 相结合的可伸缩编码方案,在较宽的比特率变化范围内改进了编码效率。

结合该课题的研究,作者已发表或录用 6 篇论文,其中属 SCI 和 EI 检索源的分别为 1 篇和 5 篇,且申请发明专利一项。该学位论文立论正确,分析深入,有创新性。表明作者在本门学科上具有坚实的理论基础和系统深入的专门知识,独立工作能力强。在答辩过程中,回答问题思路清晰,能予以正确回答。

答辩委员会表决结果

经答辩委员会表决，全票同意通过李志刚同学的博士学位论文答辩，建议授予工学博士学位。

答辩委员会主席：**余松煜**

2005年9月12日

摘　要

H.264/AVC是ITU-T和ISO联合推出的新标准，采用和发展了近几年视频编码方面的先进技术，以较高编码效率和网络友好性而著称。它仍基于以前视频编码标准的运动补偿混合编码方案，主要不同有：增强的运动预测能力；准确匹配的较小块变换；自适应环内滤波器；增强的熵编码。测试结果表明这些新特征使编码效率比以前的标准约提高50%，但增加了复杂度。本文着重研究经传输层的若干应用问题。

随着新的视频编码标准H.264/AVC的逐步推广应用，如何将H.264/AVC视频流经MPEG-2的系统传输层传输，对于充分利用原有大量的MPEG-2系统有重要意义。对此提出一种解决方案，要点是：先把H.264/AVC视频作为MPEG-2系统层传输的基本流，然后扩展MPEG-2标准中的传输流系统目标解码器(T-STD)，使之可以将H.264/AVC编码视频复用到MPEG-2传输流(TS)包中在Internet上传输和解码。被解码的基本流通常来自一个“容器”(如AVI或者TS)，在客户端从服务器端的这个容器中取出H.264/AVC基本流后便可实时解码、显示。经仿真实验表明，该方案能够获得较好的流视频效果，在带宽受限的情况下信噪比低于40dB的帧数少于5%，可用于IP网络流视频或移动视频中。

在完成H.264算法优化和多媒体指令集优化之后，将其扩展至基于比特平面编码的FGS编码，从而实现了基于PC平台

的实时 SNR(信噪比)精细粒度可伸缩编码,CIF 格式编码帧速可达 30 fps 以上。此外,为使 H. 264/AVC 码流能在带宽较大变化的 IP 和无线等异构网络中传输,本文另外提出一种将 H. 264扩展至混合空域/时域/SNR 精细可伸缩方案,并给出了相应的码率控制策略:根据率失真优化的结果来选择 QP(量化参数),在编码器端对基本层作 GOP(图像组)级的码率控制,而对增强层作逐次精细化的码率控制。本文方案与由 JM8. 6 的基本层码率控制、FGS 比特平面截断作增强层码率控制(简称 JM8. 6+FGS)的方法相比,视频质量更高,PSNR(峰值信噪比)变化更为平滑。且在接收端,在某一目标比特率约束下,可以由不同时间分辨率(帧速率)和不同空间分辨率(图像格式)配置的终端截取和实时解码。仿真结果表明,本文方法的亮度平均峰值信噪比(Y-PSNR)在 CIF 格式时优于 JM8. 6+FGS 方法达 2. 45 dB,且与目标比特率更为匹配;也比新近提出的 JVT-N020 提案,在平均 Y-PSNR 上有 0. 15 dB 的增益,而且图像质量更为平滑。

H. 264/AVC 支持 SP(同步预测)帧,允许不同质量比特流之间的高效切换,MPEG-4 支持 FGS 编码。本文提出一种将两者融合在一起的解决方案,在 JM 联合模型中实现了流切换的功能,并且加入 FGS 编码,使得传输的比特流既能适应因特网或无线网传输带宽的大跨度波动,又能灵活适应小范围的带宽变化。仿真实验结果表明:本文所提出方案的亮度 Y 分量峰值信噪比比 FGS 平均好 0. 47 dB,比流切换方法平均好0. 23 dB。且在专为移动业务而设计的 H. 264/AVC 扩展类上,在 3GPP/3GPP2 无线视频通用测试条件下,本文所提出方案的性能比单一

的FGS平均好0.76 dB,比单一的流切换方法平均好0.51 dB。

关键词 H.264, MPEG-2,网络抽象层,可伸缩编码,实时编码,流切换,视频传输,码率控制,空间可分级,时间可分级,信噪比可伸缩

此为试读,需要完整PDF请访问: www.ertongbook.com

Abstract

H. 264/AVC is a new video codec standard accepted by ITU-T and ISO, which adopts many advanced technologies and develops them in video codec since recent years, and becomes famous for its higher coding efficiency and better network friendship than others. Based on the motion-compensated hybrid coding scheme originated from previous video codec standards, H. 264/AVC has other important differences as follows: the enhanced motion prediction capability; accurate matching for small block-size transforming; adaptive in-loop deblocking filter; enhanced entropy coding. The experimental results are shown that these new characteristics promote the encoder efficiency by about 50% than the previous counterparts, but the complexity addition is inevitable. This dissertation focuses all our attentions on some applications studies throughout the transport layer in the IP networks.

With the increasingly extensive applications for the new emerging video coding standard, H. 264/AVC, it becomes more and more important that we transport H. 264/AVC video stream over MPEG - 2 system by exploiting all the existing large amount of the infrastructure of MPEG - 2 systems. We propose a solution for it. The key technologies

are as follows: Firstly we use an H. 264/AVC video as an elementary stream of MPEG - 2 system, then extend the Transport stream-System Target Decoder (T-STD) of MPEG - 2 standard, so as to pack the H. 264/AVC video elementary stream into MPEG - 2 Transport Stream(TS) to transport through Internet and decode it in the client. The decoded elementary stream should ordinarily come from a container, such as AVI or TS. We extract the H. 264/AVC video from this container to be real-time decoded and be represented in the client after the reception. The experiment results show that we get a good effect, with less than 5% frames whose PSNR<40 dB even in the case of bandwidth-constrained scenario. Running in the IP networks, the streaming solution can also be used for mobile video.

After the algorithms optimization and multi-media instruction-set optimization for the H. 264 encoder, the FGS (Fine-Granular Scalability) based on bit-plane coding is extended to realize a real-time SNR fine-granular scalable coding based on PC platform, and the frame rate can be 30 fps and over for the video with CIF format. Moreover, to enable transmission of H. 264/AVC bit-stream over IP and wireless heterogeneous networks with randomly variable bandwidths, another scheme of hybrid spatial/temporal/SNR refined scalability and its rate control strategy are given: the rate control for GOP (Group of Pictures) level is done on the base-layer at the encoder side to select the QP (Quantization Parameter) according to the RDO (Rate Distortion Optimization)

results, and the progressively refined rate control is done on the enhancement-layer. In comparison with the method of JM8.6+FGS (i.e. the base-layer rate control is based on JM8.6, and the enhancement-layer control is based on the FGS bit-plane truncation), our proposed gets better video quality with smoother PSNR (Peak Signal Noise Ratio) variation than the above. At the receiver, under certain target bit-rate constraint, the total bit-stream can be truncated and real-time decoded by the ends configured with different temporal resolutions (i.e. frame rates) and different spatial resolutions (i.e. video formats). The experimental results are shown that the Y-PSNR (PSNR for luma Y component) of ours with CIF format is 2.45 dB better than that of JM8.6+FGS, and ours matches the target bit-rate more fitly; additionally, ours is 0.15 dB better than recently proposed JVT-N020 in Y-PSNR on average, and ours gets smoother video quality than the other two schemes.

SP (Synchronization-Predictive) frame, which enables high efficiency of switching between two bitstreams with different qualities, is supported by H.264/AVC. And FGS video coding is supported by MPEG-4. This paper proposes a solution for combination between these two tools, and the Joint Model has been extended to support Stream Switching and FGS coding is further added into it so as to adapt to high bandwidth variations of Internet or Wireless networks and to low bandwidth variations flexibly for transmitted streams. Experimental results show that our proposed system outperforms

FGS by 0.47 dB and the H.264/AVC-based stream switching approach by 0.23 dB on average, respectively. And for H.264/AVC Extended Profile suitable for mobile services, our proposed system outperforms FGS by 0.76 dB and the H.264/AVC-based video stream switching approach by 0.51 dB on average, respectively, under 3GPP /3GPP2 wireless common test conditions.

Key words H.264, MPEG-2, network abstraction layer, scalable coding, real-time encoding, stream switching, video transmission, rate control, spatial scalability, temporal scalability, SNR scalability

目　录

第一章　绪论 …… 1
1.1　引言 …… 1
1.2　H.264/AVC 应用概述 …… 7
1.3　本文的结构与创新点 …… 14

第二章　H.264/AVC 基于 IP 和无线网络应用的基础 …… 16
2.1　H.264/AVC 在 IP 环境中应用的基础知识 …… 16
2.2　H.264/AVC 在无线环境中应用的基础知识 …… 30
2.3　小结 …… 35

第三章　H.264 视频流在 MPEG-2 系统层上流传输 …… 36
3.1　MPEG-2 系统层标准 …… 36
3.2　H.264 视频流在 MPEG-2 系统层上复用打包传输 …… 37
3.3　小结 …… 46

第四章　基于 H.264 的 FGS 改进方案及码率控制策略 …… 47
4.1　可伸缩编码概述 …… 47
4.2　H.264 FGS 分析及基于 PC 平台的实时 FGS 编码 …… 57
4.3　可扩大伸缩范围的空时分级 SNR-FGS 编码方法 …… 70
4.4　码率控制策略 …… 74
4.5　实验结果 …… 78
4.6　小结 …… 83

第五章　H.264－SS 与 FGS 结合的可伸缩自适应选择方法 …… 85
5.1　流切换(SS)的概念和特性 …… 85
5.2　流切换的应用场合和 SP/SI 帧的编解码方法 …… 88
5.3　基于 H.264 FGS 的流切换及其自适应选择算法 …… 95
5.4　仿真实验 …… 100
5.5　小结 …… 110

第六章　总结与展望 …… 112
参考文献 …… 114
致谢 …… 131

此为试读，需要完整PDF请访问：www.ertongbook.com